

SEQUENCE LISTING



<110> Genencor International, Inc.
 Bower, Benjamin
 Mitchinson, Colin
 Larenas, Edmund

<120> Cellulase Fusion Protein and Heterologous Cellulase Fusion Construct Encoding the Same

<130> GC832-PCT

<140> PCT/US2005/010242
 <141> 2005-03-25

<150> US 60/556,711
 <151> 2004-03-25

<160> 30

<170> PatentIn version 3.2

<210> 1
 <211> 1570
 <212> DNA
 <213> Trichoderma reesei

<400> 1

atgtatcgga	agttggccgt	catctcgccc	ttcttgccca	cagctcggtc	tcagtcggcc	60
tgcactctcc	aatcgagagac	tcacccgcct	ctgacatggc	agaaatgctc	gtctgggtggc	120
acttgcactc	aacagacagg	ctccgtggtc	atcgacgcca	actggcgctg	gactcacgct	180
acgaacagca	gcacgaactg	ctacgatggc	aacacttgg	gctcgaccct	atgtcctgac	240
aacgagacct	gcgcgaagaa	ctgctgtctg	gacgggtggc	cctacgcgtc	cacgtacgga	300
gttaccacga	gcggtaacag	cctctccatt	ggctttgtca	cccagtctgc	gcagaagaac	360
gttggcgctc	gcctttacct	tatggcgagc	gacacgacct	accaggaaatt	caccctgctt	420
ggcaacgagt	tctcttcga	tgttcatgtt	tcgcagctgc	cgtaagtgc	ttaccatgaa	480
cccctgacgt	atcttcttgc	gggctccca	ctgactggcc	aatttaagg	gcggcttgaa	540
cgtagctctc	tacttcgtgt	ccatggacgc	ggatgggtggc	gtgagcaagt	atcccaccaa	600
caccgctggc	gccaagta	gcaacgggta	ctgtgacagc	cagtgtcccc	gcgatctgaa	660
gttcatcaat	ggccaggcca	acgttgaggg	ctgggagcc	tcatccaaca	acgcaaacac	720
gggcatttgg	ggacacggaa	gctgctgctc	tgagatggat	atctgggagg	ccaactccat	780
ctccgaggct	tttacccccc	acccttgac	gactgtcgcc	caggagatct	gcgagggtga	840
tgggtcgcc	ggaacttact	ccgataacag	atatggcgcc	acttgcgatc	ccgatggctg	900
cgacttggaa	ccataccgccc	tggcaacac	cagcttctac	ggccctggct	caagcttac	960
cctcgatacc	accaagaaat	tgaccgtgt	cacccagttc	gagacgtcg	gtgccatcaa	1020
ccgataactat	gtccagaatg	gcgtacttt	ccagcagccc	aacgcccggc	ttggtagtta	1080
ctctggcaac	gagctcaacg	atgattactg	cacagctgag	gaggcagaat	tcggcgatc	1140
ctctttctca	gacaaggcg	gcctgactca	gttcaagaag	gctacctctg	gcggcatggt	1200
tctgggtatg	agtctgtgg	atgatgtgag	tttcatggac	aaacatgcgc	gttgacaaag	1260
agtcaaggcg	ctgactgaga	tgttacagta	ctacgccaac	atgctgtggc	tggactccac	1320
ctacccgaca	aacgagacct	cctccacacc	cggtggccgt	cgcgaaagct	gctccaccag	1380
ctccgggtgc	cctgctcagg	tcgaatctca	gtctcccaac	gccaagggtca	ccttctccaa	1440
catcaagttc	ggacccattg	gcagcaccgg	caaccctagc	ggcgcaacc	ctccccggcg	1500
aaacccgcct	ggcaccacca	ccacccggcc	cccagccact	accactggaa	gctctcccg	1560
acctactagt						1570

<210> 2

<211> 51
 <212> DNA
 <213> Trichoderma reesei

<400> 2
 atgtatcgga agttggccgt catctcgccc ttcttggcca cagctcgtgc t 51

<210> 3
 <211> 1438
 <212> DNA
 <213> Trichoderma reesei

<400> 3
 cagtcggcct gcactctcca atcggagact caccgcctc tgacatggca gaaatgctcg 60
 tctggtggca cttgcactca acagacaggc tccgtggta tcgacgcctaa ctggcgctgg 120
 actcacgcta cgaacagcag cacgaactgc tacgatggca acacttggag ctcgacccta 180
 tgtcctgaca acgagacctg cgcaagaac tgctgtctgg acgggtccgc ctacgcgtcc 240
 acgtacggag ttaccacggag cgtaacagc ctctccattt gctttgtcac ccagtctgcg 300
 cagaagaacg ttggcgctcg cctttacattt atggcgagcg acacgaccta ccaggaattc 360
 accctgcttgc gcaacgagtt ctcttcgat gttgatgttt cgcaagtcgtcc gtaagtgtact 420
 taccatgaac ccctgacgta tcttcttgc ggctcccagc tgactggcca atttaaggtg 480
 cggcttgaac ggagctctt acttcgtgtc catggacgcg gatgggtggcg tgagcaagta 540
 tcccaccaac accgctggcg ccaagtaacgg cacgggggtac tggacagcc agtgcgtcc 600
 cgatctgaag ttcatcaatg gccaggccaa cggtgaggc tgggagccgt catccaacaa 660
 cgcaaaacacg ggcattggag gacacggaaag ctgctgtct gagatggata tctgggaggc 720
 caactccatc tccgaggctc ttacccca cccttgcacg actgtcggcc aggagatctg 780
 cgagggtgat ggggtgcggcg gaacttaactc cgataacaga tatggcgcc cttgcgtatcc 840
 cgatggctgc gactggaaacc cataccgcct gggcaacacc agcttctacg gcccctggctc 900
 aagctttacc ctgcatacca ccaagaaatt gaccgttgc acccagttcg agacgtcg 960
 tgccatcaac cgataactatg tccagaatgg cgtcacttgc cagcagccca acgcccagct 1020
 tggtagttac tctggcaacg agctcaacga tgattactgc acagctgagg aggcagaatt 1080
 cggcgatcc tctttctcag acaagggcggt cctgactcgttcaagaagg ctacccctgg 1140
 cggcatgggt ctggcatga gtctgtggta tgatgtgagt ttgatggaca aacatgcgcg 1200
 ttgacaaaga gtcaagcagc tgactgagat gttacagttac tacgccaaca tgctgtggct 1260
 ggactccacc taccgcacaa acgagacccctc ctccacaccc ggtgcccgtgc gcgaaagctg 1320
 ctccaccaggc tccgggtgtcc ctgctcaggat cgaatctcgt tctcccaacg ccaagggtcac 1380
 cttctccaac atcaagttcg gaccattgg cagcaccggc aaccctagcg gcgcaac 1438

<210> 4
 <211> 81
 <212> DNA
 <213> Trichoderma reesei

<400> 4
 cctcccgccg gaaacccgccc tgccaccacc accacccgccc gcccagccac taccactgg 60
 agctctcccg gacctactag t 81

<210> 5
 <211> 480
 <212> PRT
 <213> Trichoderma reesei

<400> 5
 Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg
 1 5 10 15
 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr

20	25	30	
Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser			
35	40	45	
Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser			
50	55	60	
Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp			
65	70	75	80
Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala			
85	90	95	
Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe			
100	105	110	
Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met			
115	120	125	
Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe			
130	135	140	
Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala			
145	150	155	160
Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro			
165	170	175	
Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln			
180	185	190	
Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly			
195	200	205	
Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly			
210	215	220	
Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu			
225	230	235	240
Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu			
245	250	255	
Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr			
260	265	270	
Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr			
275	280	285	
Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys			
290	295	300	
Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr			
305	310	315	320
Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly			
325	330	335	
Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu			
340	345	350	
Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln			
355	360	365	
Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp			
370	375	380	
Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr			
385	390	395	400
Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr			
405	410	415	
Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys			
420	425	430	
Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn			
435	440	445	
Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr			
450	455	460	
Thr Arg Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Ser			
465	470	475	480

<210> 6
<211> 431
<212> PRT
<213> Trichoderma reesei

<400> 6
Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
1 5 10 15
Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
20 25 30
Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
35 40 45
Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
50 55 60
Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
65 70 75 80
Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
85 90 95
Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
100 105 110
Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
115 120 125
Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
130 135 140
Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
145 150 155 160
Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
165 170 175
Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
180 185 190
Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
195 200 205
Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
210 215 220
Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
225 230 235 240
Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
245 250 255
Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
260 265 270
Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
275 280 285
Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
290 295 300
Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
305 310 315 320
Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
325 330 335
Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
340 345 350
Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
355 360 365
Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
370 375 380
Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
385 390 395 400

Ser	Gly	Val	Pro	Ala	Gln	Val	Glu	Ser	Gln	Ser	Pro	Asn	Ala	Lys	Val
405								410						415	
Thr	Phe	Ser	Asn	Ile	Lys	Phe	Gly	Pro	Ile	Gly	Ser	Thr	Gly	Asn	
420								425						430	

<210> 7
 <211> 1077
 <212> DNA
 <213> Acidothermus cellulolyticus

<400> 7

gcgggcggcg	gctattggca	cacgagcggc	cgggagatcc	tggacgcgaa	caacgtgccg	60
gtacggatcg	ccggcatcaa	ctggtttggg	ttcgaaacct	gcaattacgt	cgtgcacgg	120
ctctggtcac	gcgactaccg	cagcatgctc	gaccagataa	agtgcgtcgg	ctacaacaca	180
atccggctgc	cgtactctga	cgacattctc	aagccgggca	ccatgcccga	cagcatcaat	240
ttttaccaga	tgaatcagga	cctgcagggt	ctgacgtcct	tgcaggtcat	ggacaaaatc	300
gtcgcgtacg	ccggtcagat	cggcctgcgc	atcattctt	accgccaccg	accggattgc	360
agcggggcagt	ccggcgtgt	gtacacgagc	agcgtctcgg	aggctacgt	gatttccgac	420
ctgcaagcgc	tggcgcagcg	ctacaaggga	aacccgacgg	tgcgtggctt	tgacttgcac	480
aacgagccgc	atgaccggc	ctgctggggc	tgcggcgatc	cgagcatcga	ctggcgattg	540
gccgcccggc	gggcccggaaa	cgccgtgctc	tcggtaatc	cgaacctgct	catttcgtc	600
gaaggtgtc	agagctacaa	cgagactcc	tactggtggg	gcccgaacct	gcaaggagcc	660
ggccagtacc	cggtcgtgct	gaacgtgccc	aaccgcctgg	tgtactcggc	gcacgactac	720
gacgacgagcg	tctacccca	gacgtggttc	agcgatccga	ccttccccaa	caacatgccc	780
ggcatctgga	acaagaactg	gggatacctc	ttcaatcaga	acattgcacc	ggtatggctg	840
ggcgaattcg	gtacgacact	gcaatccacg	accgaccaga	cgtggctgaa	gacgctcgac	900
cagtacctac	ggccgaccgc	gcaatacgg	gcccggacagct	tccagtggac	cttctggtcc	960
tggaaccccg	attccggcga	cacaggagga	attctcaagg	atgactggca	gacggctcgac	1020
acagtaaaag	acggctatct	cgcggcgatc	aagtcgtcga	ttttcgatcc	tgtcggc	1077

<210> 8
 <211> 359
 <212> PRT
 <213> Acidothermus cellulolyticus

<400> 8

Ala	Gly	Gly	Gly	Tyr	Trp	His	Thr	Ser	Gly	Arg	Glu	Ile	Leu	Asp	Ala
1				5			10				15				
Asn	Asn	Val	Pro	Val	Arg	Ile	Ala	Gly	Ile	Asn	Trp	Phe	Gly	Phe	Glu
						20			25			30			
Thr	Cys	Asn	Tyr	Val	Val	His	Gly	Leu	Trp	Ser	Arg	Asp	Tyr	Arg	Ser
						35			40			45			
Met	Leu	Asp	Gln	Ile	Lys	Ser	Leu	Gly	Tyr	Asn	Thr	Ile	Arg	Leu	Pro
						50			55			60			
Tyr	Ser	Asp	Asp	Ile	Leu	Lys	Pro	Gly	Thr	Met	Pro	Asn	Ser	Ile	Asn
						65			70			75			80
Phe	Tyr	Gln	Met	Asn	Gln	Asp	Leu	Gln	Gly	Leu	Thr	Ser	Leu	Gln	Val
						85			90			95			
Met	Asp	Lys	Ile	Val	Ala	Tyr	Ala	Gly	Gln	Ile	Gly	Leu	Arg	Ile	Ile
						100			105			110			
Leu	Asp	Arg	His	Arg	Pro	Asp	Cys	Ser	Gly	Gln	Ser	Ala	Leu	Trp	Tyr
						115			120			125			
Thr	Ser	Ser	Val	Ser	Glu	Ala	Thr	Trp	Ile	Ser	Asp	Leu	Gln	Ala	Leu
						130			135			140			
Ala	Gln	Arg	Tyr	Lys	Gly	Asn	Pro	Thr	Val	Val	Gly	Phe	Asp	Leu	His
						145			150			155			160
Asn	Glu	Pro	His	Asp	Pro	Ala	Cys	Trp	Gly	Cys	Gly	Asp	Pro	Ser	Ile

```
<210> 9
<211> 1914
<212> DNA
<213> Acidothermus cellulolyticus
```

<400>	9								
aacgaccgg	atcatccagcg	gttcctc	acg	atgtacaaca	agattcacga	cccagcgaac	60		
ggctacttca	gcccgcagg	aattccctac	cactcggtag	aaacgctcat	cgttgaggca	120			
ccggactacg	ggcacgagac	aacttcggag	gcgtacagct	tctggctctg	gctcgaagcg	180			
acgtacggcg	cagtgaccgg	caacttgacg	ccgttcaaca	acgcctggac	gacgatggaa	240			
acgtacatga	tcccgcagca	cgcggaccag	ccgaacaacg	cgtcgta	ccccaacacg	300			
ccggcgtcgt	acgctccgg	agagccg	ctg	cccagcatgt	accgg	360			
agcgtgccgg	ttgggcacga	cccgc	tcg	cc	gcaattgc	agtgcacgt	cggcactccg	420	
gacatttacg	gcatgcactg	gctggccgac	gtt	gacaaca	tctacggata	cggcgac	480		
cccgccgg	gttgcgaact	cgg	tc	ttcc	gtaagg	ggcg	540		
cagcgcgg	cgcaggagtc	cgt	ctggg	gag	acggtcac	cc	agccgac	600	
aagtacgg	gggcgcacgg	ctacgtc	gac	ctgtt	catcc	agg	gttcgac	660	
tggaagtaca	ccgatgcccc	ggacggc	gac	gccc	gtgc	tcc	agg	gtc	720
tacac	ctcg	ggc	caagg	ca	agcgc	gatt	ccc	gac	780
gcca	actcg	gcg	actac	ctcg	ctctt	gaca	ag	actt	840
aact	gttac	cg	ccag	ctc	ctg	cc	gg	ttc	900
atcgg	ctggt	actac	gc	tc	gg	cc	gg	ct	960
ggc	ccgc	actt	cgg	ct	cc	ag	gg	at	1020
ccg	ctt	cgt	tc	cg	cc	g	gg	at	1080
ctgg	gagg	tt	t	cc	at	cc	gg	cc	1140
agct	ggaa	ac	g	cc	at	cc	gg	cc	1200
tacg	actt	gg	gg	ct	cc	ac	cc	gg	1260
tgg	ttccat	gg	aa	c	cc	gt	cc	gg	1320
ctg	ctcg	aca	at	gg	gg	tc	cc	ac	1380
at	ccgt	ca	tt	gg	gg	tc	ac	cc	1440
acga	aatg	cc	ac	ct	gc	ac	gt	cc	1500

gcgctcgca agacactcg a gtactacgca gcaaaatccg gcgatacggc ctcgcgcgac	1560
ctcgcaagg gattgctcg a ctccatgtgg aacaacgacc aggacagccg cgggtgtgagc	1620
acaccggaga cgccgaccg a ctactctcg ttcactcagg tgtacgacc gacgactgg	1680
gacggcctct acatcccgtc gggttggacg gggaccatgc ccaacggta ccaaatacg	1740
ccgggtgcga ctttcctgag catccggtcc tggtacacca aggatccgca gtggcgaag	1800
gtcaggcgt acctcaacgg cggcctgct ccgacgttca actaccaccc gttctggcg	1860
gagtccgact tcgcgatggc gaacgcccgt tttggcatgc tcttccatc cggg	1914

<210> 10
 <211> 638
 <212> PRT
 <213> Acidothermus cellulolyticus

<400> 10	
Asn Asp Pro Tyr Ile Gln Arg Phe Leu Thr Met Tyr Asn Lys Ile His	
1 5 10 15	
Asp Pro Ala Asn Gly Tyr Phe Ser Pro Gln Gly Ile Pro Tyr His Ser	
20 25 30	
Val Glu Thr Leu Ile Val Glu Ala Pro Asp Tyr Gly His Glu Thr Thr	
35 40 45	
Ser Glu Ala Tyr Ser Phe Trp Leu Trp Leu Glu Ala Thr Tyr Gly Ala	
50 55 60	
Val Thr Gly Asn Trp Thr Pro Phe Asn Asn Ala Trp Thr Thr Met Glu	
65 70 75 80	
Thr Tyr Met Ile Pro Gln His Ala Asp Gln Pro Asn Asn Ala Ser Tyr	
85 90 95	
Asn Pro Asn Ser Pro Ala Ser Tyr Ala Pro Glu Glu Pro Leu Pro Ser	
100 105 110	
Met Tyr Pro Val Ala Ile Asp Ser Ser Val Pro Val Gly His Asp Pro	
115 120 125	
Leu Ala Ala Glu Leu Gln Ser Thr Tyr Gly Thr Pro Asp Ile Tyr Gly	
130 135 140	
Met His Trp Leu Ala Asp Val Asp Asn Ile Tyr Gly Tyr Gly Asp Ser	
145 150 155 160	
Pro Gly Gly Cys Glu Leu Gly Pro Ser Ala Lys Gly Val Ser Tyr	
165 170 175	
Ile Asn Thr Phe Gln Arg Gly Ser Gln Glu Ser Val Trp Glu Thr Val	
180 185 190	
Thr Gln Pro Thr Cys Asp Asn Gly Lys Tyr Gly Gly Ala His Gly Tyr	
195 200 205	
Val Asp Leu Phe Ile Gln Gly Ser Thr Pro Pro Gln Trp Lys Tyr Thr	
210 215 220	
Asp Ala Pro Asp Ala Asp Ala Arg Ala Val Gln Ala Ala Tyr Trp Ala	
225 230 235 240	
Tyr Thr Trp Ala Ser Ala Gln Gly Lys Ala Ser Ala Ile Ala Pro Thr	
245 250 255	
Ile Ala Lys Ala Ala Lys Leu Gly Asp Tyr Leu Arg Tyr Ser Leu Phe	
260 265 270	
Asp Lys Tyr Phe Lys Gln Val Gly Asn Cys Tyr Pro Ala Ser Ser Cys	
275 280 285	
Pro Gly Ala Thr Gly Arg Gln Ser Glu Thr Tyr Leu Ile Gly Trp Tyr	
290 295 300	
Tyr Ala Trp Gly Gly Ser Ser Gln Gly Trp Ala Trp Arg Ile Gly Asp	
305 310 315 320	
Gly Ala Ala His Phe Gly Tyr Gln Asn Pro Leu Ala Ala Trp Ala Met	
325 330 335	
Ser Asn Val Thr Pro Leu Ile Pro Leu Ser Pro Thr Ala Lys Ser Asp	

340	345	350	
Trp Ala Ala Ser Leu Gln Arg Gln	Leu Glu Phe Tyr Gln	Trp Leu Gln	
355	360	365	
Ser Ala Glu Gly Ala Ile Ala Gly	Gly Ala Thr Asn Ser Trp Asn Gly		
370	375	380	
Asn Tyr Gly Thr Pro Pro Ala Gly	Asp Ser Thr Phe Tyr Gly Met Ala		
385	390	395	400
Tyr Asp Trp Glu Pro Val Tyr His	Asp Pro Pro Ser Asn Asn Trp Phe		
405	410	415	
Gly Phe Gln Ala Trp Ser Met Glu	Arg Val Ala Glu Tyr Tyr Val		
420	425	430	
Thr Gly Asp Pro Lys Ala Lys Ala	Leu Leu Asp Lys Trp Val Ala Trp		
435	440	445	
Val Lys Pro Asn Val Thr Thr Gly	Ala Ser Trp Ser Ile Pro Ser Asn		
450	455	460	
Leu Ser Trp Ser Gly Gln Pro Asp	Thr Trp Asn Pro Ser Asn Pro Gly		
465	470	475	480
Thr Asn Ala Asn Leu His Val	Thr Ile Thr Ser Ser Gly Gln Asp	Val	
485	490	495	
Gly Val Ala Ala Ala Leu Ala Lys	Thr Leu Glu Tyr Tyr Ala Ala Lys		
500	505	510	
Ser Gly Asp Thr Ala Ser Arg Asp	Leu Ala Lys Gly Leu Leu Asp Ser		
515	520	525	
Met Trp Asn Asn Asp Gln Asp Ser	Leu Gly Val Ser Thr Pro Glu Thr		
530	535	540	
Arg Thr Asp Tyr Ser Arg Phe	Thr Gln Val Tyr Asp Pro Thr Thr Gly		
545	550	555	560
Asp Gly Leu Tyr Ile Pro Ser Gly	Trp Thr Gly Thr Met Pro Asn Gly		
565	570	575	
Asp Gln Ile Lys Pro Gly Ala Thr	Phe Leu Ser Ile Arg Ser Trp Tyr		
580	585	590	
Thr Lys Asp Pro Gln Trp Ser Lys	Val Gln Ala Tyr Leu Asn Gly Gly		
595	600	605	
Pro Ala Pro Thr Phe Asn Tyr His	Arg Phe Trp Ala Glu Ser Asp Phe		
610	615	620	
Ala Met Ala Asn Ala Asp Phe Gly	Met Leu Phe Pro Ser Gly		
625	630	635	

<210> 11
 <211> 2223
 <212> DNA
 <213> Acidothermus cellulolyticus

<400> 11

gcgcacgactc	agccgtacac	ctggagcaac	gtggcgatcg	ggggcggcgg	ctttgtcgac	60
gggatcgct	tcaatgaagg	tgcacccggga	attctgtacg	tgccgacgga	catcgggggg	120
atgtatcgat	gggatgccgc	caacgggcgg	tggatccctc	ttctggattt	ggtgggatgg	180
aacaatttggg	ggtacaacgg	cgtcgtcagc	attgcggcag	acccgatcaa	tactaacaag	240
gtatgggccc	ccgtcggaat	gtacaccaac	agctgggacc	caaacgcacgg	agcgattctc	300
cgtcgctctg	atcagggcgc	aacgtggcaa	ataacgcccc	tgccgttcaa	gcttggcggc	360
aacatgccc	ggcgttggat	gggcgagcgg	cttgcgggtgg	atccaaaccaa	tgacaacatt	420
ctgtatcc	gcgc	ccccgag	cggcaaagggg	ctctggagaa	gcacagattc	480
tggtcccaga	tgacgaactt	tccggacgtta	ggcacgtaca	ttgcaaattcc	cactgacacg	540
accggctata	agagcgatata	tcaaggcg	gtctgggtcg	ctttcgacaa	gtcttcgtca	600
tcgctcgggc	aagcgagtaa	gaccat	tttggcgtgg	cgatcccaa	taatccggtc	660
ttctggagca	gagacggcgg	cgcgacgtgg	caggcgggtgc	cggtgcgc	gaccggcttc	720
atcccgaca	agggcg	tct	tgacccggtc	aaccacgtgc	tctatattgc	780

acgggtggtc	cgtatgacgg	gagctccggc	gacgtctgga	aattctcggt	gacctccggg	840
acatggacgc	gaatcagccc	ggtaccttcg	acggacacgg	ccaacgacta	ctttggttac	900
agcggcctca	ctatcgacccg	ccagcaccccg	aacacgataa	tggtggcaac	ccagatatcg	960
tggtggccgg	acaccataat	cttcggagc	accgacggcg	gtgcgacgtg	gacgcggatc	1020
tgggatttga	cgagttatcc	caatcgaaagc	ttgcgatatg	tgcttgacat	ttcggcggag	1080
ccttggctga	ccttcggcgt	acagccgaat	cctcccgatc	cgagtccgaa	gctcggctgg	1140
atggatgaag	cgatggcaat	cgatccgttc	aactctgatc	ggatgctcta	cggaacagggc	1200
gchgacgttgt	acgcaacaaa	tgcacacg	aagtgggact	ccggcggcca	gattcatatc	1260
gchccgatgg	tcaaggatt	ggaggagacg	gchgttaaacg	atctcatcg	ccgcccgtct	1320
ggcgccccgc	tcatcagcgc	tctcgagac	ctcggcggt	tcacccacgc	cgacgttact	1380
gcccgtccat	cgacgatctt	cacgtcaccg	gtttcacga	ccggcaccag	cgtcgactat	1440
gchgaattga	atccgtcgat	catcggtcgc	gctggaaattt	tcgatccatc	gagccaaaccg	1500
aacgacagggc	acgtcgctt	ctcgacagac	ggcggcaaga	actggttcca	aggcagcgaa	1560
cctggcgggg	tgacgacggg	cgccaccgtc	gccgcatcgg	ccgacggctc	tcgttcgtc	1620
tgggctcccg	gcgatcccg	tcagcctgtg	gtgtacgcag	tcggatttgg	caactcctgg	1680
gctgcttcgc	aagggtttcc	cgccaatgccc	cagatccgct	cagaccgggt	gaatccaaag	1740
actttctatg	ccctatccaa	tggAACCTTC	tatcgaaagca	cggaacggcg	cgtgacattc	1800
caaccggctg	cgccgggtct	tccgagcagc	ggtgcgtcg	gtgtcatgtt	ccacgcgggt	1860
cctggaaaag	aaggcgatct	gtggctcgct	gcatcgagcg	ggctttacca	ctcaaccaat	1920
ggcggcagca	gttggctcgc	aatcaccggc	gtatccctcg	cggtgaacgt	gggatttgg	1980
aagtctgcgc	ccgggtcgct	atacccagcc	gtctttgtcg	tcggcacgat	cggaggcggt	2040
acggggcggt	accgctccga	cgacgggtggg	acgacctggg	tacggatcaa	tgcgtaccag	2100
caccaatacg	gaaattgggg	acaagcaatc	accggtgacc	cgcgaattta	cgggcgggtg	2160
tacataggca	cgaacggccg	tggattgtc	tacggggaca	ttgggtgtc	gccgtccgga	2220
						2223

<210> 12
 <211> 741
 <212> PRT
 <213> Acidothermus cellulolyticus

<400> 12
 Ala Thr Thr Gln Pro Tyr Thr Trp Ser Asn Val Ala Ile Gly Gly Gly
 1 5 10 15
 Gly Phe Val Asp Gly Ile Val Phe Asn Glu Gly Ala Pro Gly Ile Leu
 20 25 30
 Tyr Val Arg Thr Asp Ile Gly Gly Met Tyr Arg Trp Asp Ala Ala Asn
 35 40 45
 Gly Arg Trp Ile Pro Leu Leu Asp Trp Val Gly Trp Asn Asn Trp Gly
 50 55 60
 Tyr Asn Gly Val Val Ser Ile Ala Ala Asp Pro Ile Asn Thr Asn Lys
 65 70 75 80
 Val Trp Ala Ala Val Gly Met Tyr Thr Asn Ser Trp Asp Pro Asn Asp
 85 90 95
 Gly Ala Ile Leu Arg Ser Ser Asp Gln Gly Ala Thr Trp Gln Ile Thr
 100 105 110
 Pro Leu Pro Phe Lys Leu Gly Gly Asn Met Pro Gly Arg Gly Met Gly
 115 120 125
 Glu Arg Leu Ala Val Asp Pro Asn Asn Asp Asn Ile Leu Tyr Phe Gly
 130 135 140
 Ala Pro Ser Gly Lys Gly Leu Trp Arg Ser Thr Asp Ser Gly Ala Thr
 145 150 155 160
 Trp Ser Gln Met Thr Asn Phe Pro Asp Val Gly Thr Tyr Ile Ala Asn
 165 170 175
 Pro Thr Asp Thr Thr Gly Tyr Gln Ser Asp Ile Gln Gly Val Val Trp
 180 185 190
 Val Ala Phe Asp Lys Ser Ser Ser Leu Gly Gln Ala Ser Lys Thr

195	200	205
Ile Phe Val Gly Val Ala Asp Pro Asn Asn Pro Val Phe Trp Ser Arg		
210	215	220
Asp Gly Gly Ala Thr Trp Gln Ala Val Pro Gly Ala Pro Thr Gly Phe		
225	230	235
Ile Pro His Lys Gly Val Phe Asp Pro Val Asn His Val Leu Tyr Ile		
245	250	255
Ala Thr Ser Asn Thr Gly Gly Pro Tyr Asp Gly Ser Ser Gly Asp Val		
260	265	270
Trp Lys Phe Ser Val Thr Ser Gly Thr Trp Thr Arg Ile Ser Pro Val		
275	280	285
Pro Ser Thr Asp Thr Ala Asn Asp Tyr Phe Gly Tyr Ser Gly Leu Thr		
290	295	300
Ile Asp Arg Gln His Pro Asn Thr Ile Met Val Ala Thr Gln Ile Ser		
305	310	315
Trp Trp Pro Asp Thr Ile Ile Phe Arg Ser Thr Asp Gly Gly Ala Thr		
325	330	335
Trp Thr Arg Ile Trp Asp Trp Thr Ser Tyr Pro Asn Arg Ser Leu Arg		
340	345	350
Tyr Val Leu Asp Ile Ser Ala Glu Pro Trp Leu Thr Phe Gly Val Gln		
355	360	365
Pro Asn Pro Pro Val Pro Ser Pro Lys Leu Gly Trp Met Asp Glu Ala		
370	375	380
Met Ala Ile Asp Pro Phe Asn Ser Asp Arg Met Leu Tyr Gly Thr Gly		
385	390	395
Ala Thr Leu Tyr Ala Thr Asn Asp Leu Thr Lys Trp Asp Ser Gly Gly		
405	410	415
Gln Ile His Ile Ala Pro Met Val Lys Gly Leu Glu Glu Thr Ala Val		
420	425	430
Asn Asp Leu Ile Ser Pro Pro Ser Gly Ala Pro Leu Ile Ser Ala Leu		
435	440	445
Gly Asp Leu Gly Gly Phe Thr His Ala Asp Val Thr Ala Val Pro Ser		
450	455	460
Thr Ile Phe Thr Ser Pro Val Phe Thr Thr Gly Thr Ser Val Asp Tyr		
465	470	475
Ala Glu Leu Asn Pro Ser Ile Ile Val Arg Ala Gly Ser Phe Asp Pro		
485	490	495
Ser Ser Gln Pro Asn Asp Arg His Val Ala Phe Ser Thr Asp Gly Gly		
500	505	510
Lys Asn Trp Phe Gln Gly Ser Glu Pro Gly Gly Val Thr Thr Gly Gly		
515	520	525
Thr Val Ala Ala Ser Ala Asp Gly Ser Arg Phe Val Trp Ala Pro Gly		
530	535	540
Asp Pro Gly Gln Pro Val Val Tyr Ala Val Gly Phe Gly Asn Ser Trp		
545	550	555
Ala Ala Ser Gln Gly Val Pro Ala Asn Ala Gln Ile Arg Ser Asp Arg		
565	570	575
Val Asn Pro Lys Thr Phe Tyr Ala Leu Ser Asn Gly Thr Phe Tyr Arg		
580	585	590
Ser Thr Asp Gly Gly Val Thr Phe Gln Pro Val Ala Ala Gly Leu Pro		
595	600	605
Ser Ser Gly Ala Val Gly Val Met Phe His Ala Val Pro Gly Lys Glu		
610	615	620
Gly Asp Leu Trp Leu Ala Ala Ser Ser Gly Leu Tyr His Ser Thr Asn		
625	630	635
Gly Gly Ser Ser Trp Ser Ala Ile Thr Gly Val Ser Ser Ala Val Asn		
645	650	655

Val Gly Phe Gly Lys Ser Ala Pro Gly Ser Ser Tyr Pro Ala Val Phe
 660 665 670
 Val Val Gly Thr Ile Gly Gly Val Thr Gly Ala Tyr Arg Ser Asp Asp
 675 680 685
 Gly Gly Thr Thr Trp Val Arg Ile Asn Asp Asp Gln His Gln Tyr Gly
 690 695 700
 Asn Trp Gly Gln Ala Ile Thr Gly Asp Pro Arg Ile Tyr Gly Arg Val
 705 710 715 720
 Tyr Ile Gly Thr Asn Gly Arg Gly Ile Val Tyr Gly Asp Ile Gly Gly
 725 730 735
 Ala Pro Ser Gly Ser
 740

<210> 13
 <211> 1677
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> construct based on *Thermobifida fusca*

<400> 13
 gccggctgct cgggtggacta cacggtaaac tcctgggta ccgggttcac cgccaaacgtc 60
 accatcacca acctcgccag tgcgtatcaac ggctggaccc tggagtggga cttcccccggc 120
 aaccagcagg tgaccaaacct gtggaaacggg acctacaccc agtccgggca gcacgtgtcg 180
 gtcagcaacg ccccgtaaca cgcctccatc ccggccaaacg gaacgggttga gttcgggttc 240
 aacggctcct actcgccgag caacgacatc ccctcctcct tcaagctgaa cgggggttacc 300
 tgcgacggct cggacgaccc cgaccccgag cccagccccct ccccccagccc ttccccccagc 360
 cccacagacc cggatgagcc gggcgccccg accaaccgc ccaccaaccc cggcgagaag 420
 gtcgacaacc cgttcgaggg cgccaaagctg tacgtgaacc cggctctggc gccaaggcc 480
 gccgctgagc cgggcgggttc cgggtcgcc aacgagtcca ccgctgtctg gctggaccgt 540
 atcggcgcca tcgagggcaa cgacagccccg accaccggct ccatgggtct gcgcgaccac 600
 ctggaggagg cctgtccgcca gtccgggtggc gacccgctga ccatccaggc cgtcatctac 660
 aacctgcccc gccgcgactg cggccgcgtg gcctccaacg gtgagctggg tcccgatgaa 720
 ctcgaccgct acaagagcga gtacatcgac ccgatcgccg acatcatgtg ggacttcgca 780
 gactacgaga acctcgccat cgtcgccatc atcgagatcg actccctgcc caacccgtc 840
 accaacgtgg gcggaacacgg cggcaccgag ctctgcgcct acatgaagca gAACGGCGGC 900
 tacgtcaacg gtgtcggtca cggccctccgc aagctggcg agatcccgaa cgtctacaac 960
 tacatcgacg cggccacca cggctggatc ggctggact ccaacttcgg cccctcggtg 1020
 gacatcttct acgaggccgc caacgcctcc ggctccaccg tggactacgt gcacggcttc 1080
 atctccaaca cggccaacta ctcggccact gtggagccgt acctggacgt caacggcacc 1140
 gttAACGGCC agctcatccg ccagtccaaag tgggttact ggaaccagta cgtcgacgag 1200
 ctctccttcg tccaggacat cgtcgccatc ctgatcgcca agggcttcgg gtccgacatc 1260
 ggtatgctca tcgacaccc cccgtccatc cggcaacggc tgggttggcc cgaaccgtcc gaccggaccg 1320
 agctccctca cccgtccatc cacctacgtt gacgagagcc gtatcgaccc cccgtatccac 1380
 cccggtaact ggtgcaacca gggccgtgcg ggcctcgcc agcggccac ggtcaacccg 1440
 gctcccggtg ttgacgccta cgtctgggtg aagccccccgg gtgagtcgaa cggcgccacg 1500
 gaggagatcc cgaacgacga gggcaaggcc ttcgaccgca tgcgtgcaccc gacctaccag 1560
 ggcaacggccc gcaacggcaa caacccctcg ggtgcgtgc ccaacgcccc catctccggc 1620
 cactggttct ctggccagtt cggcgagctg ctggccaaacg cctacccggc tctgtaa 1677

<210> 14
 <211> 558
 <212> PRT
 <213> Artificial Sequence

<220>

<223> construct based on *Thermobifida fusca*

<400> 14
Ala Gly Cys Ser Val Asp Tyr Thr Val Asn Ser Trp Gly Thr Gly Phe
1 5 10 15
Thr Ala Asn Val Thr Ile Thr Asn Leu Gly Ser Ala Ile Asn Gly Trp
20 25 30
Thr Leu Glu Trp Asp Phe Pro Gly Asn Gln Gln Val Thr Asn Leu Trp
35 40 45
Asn Gly Thr Tyr Thr Gln Ser Gly Gln His Val Ser Val Ser Asn Ala
50 55 60
Pro Tyr Asn Ala Ser Ile Pro Ala Asn Gly Thr Val Glu Phe Gly Phe
65 70 75 80
Asn Gly Ser Tyr Ser Gly Ser Asn Asp Ile Pro Ser Ser Phe Lys Leu
85 90 95
Asn Gly Val Thr Cys Asp Gly Ser Asp Asp Pro Asp Pro Glu Pro Ser
100 105 110
Pro Ser Pro Ser Pro Ser Pro Ser Pro Thr Asp Pro Asp Glu Pro Gly
115 120 125
Gly Pro Thr Asn Pro Pro Thr Asn Pro Gly Glu Lys Val Asp Asn Pro
130 135 140
Phe Glu Gly Ala Lys Leu Tyr Val Asn Pro Val Trp Ser Ala Lys Ala
145 150 155 160
Ala Ala Glu Pro Gly Gly Ser Ala Val Ala Asn Glu Ser Thr Ala Val
165 170 175
Trp Leu Asp Arg Ile Gly Ala Ile Glu Gly Asn Asp Ser Pro Thr Thr
180 185 190
Gly Ser Met Gly Leu Arg Asp His Leu Glu Glu Ala Val Arg Gln Ser
195 200 205
Gly Gly Asp Pro Leu Thr Ile Gln Val Val Ile Tyr Asn Leu Pro Gly
210 215 220
Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Leu Gly Pro Asp Glu
225 230 235 240
Leu Asp Arg Tyr Lys Ser Glu Tyr Ile Asp Pro Ile Ala Asp Ile Met
245 250 255
Trp Asp Phe Ala Asp Tyr Glu Asn Leu Arg Ile Val Ala Ile Ile Glu
260 265 270
Ile Asp Ser Leu Pro Asn Leu Val Thr Asn Val Gly Gly Asn Gly Gly
275 280 285
Thr Glu Leu Cys Ala Tyr Met Lys Gln Asn Gly Gly Tyr Val Asn Gly
290 295 300
Val Gly Tyr Ala Leu Arg Lys Leu Gly Glu Ile Pro Asn Val Tyr Asn
305 310 315 320
Tyr Ile Asp Ala Ala His His Gly Trp Ile Gly Trp Asp Ser Asn Phe
325 330 335
Gly Pro Ser Val Asp Ile Phe Tyr Glu Ala Ala Asn Ala Ser Gly Ser
340 345 350
Thr Val Asp Tyr Val His Gly Phe Ile Ser Asn Thr Ala Asn Tyr Ser
355 360 365
Ala Thr Val Glu Pro Tyr Leu Asp Val Asn Gly Thr Val Asn Gly Gln
370 375 380
Leu Ile Arg Gln Ser Lys Trp Val Asp Trp Asn Gln Tyr Val Asp Glu
385 390 395 400
Leu Ser Phe Val Gln Asp Leu Arg Gln Ala Leu Ile Ala Lys Gly Phe
405 410 415
Arg Ser Asp Ile Gly Met Leu Ile Asp Thr Ser Arg Asn Gly Trp Gly
420 425 430

Gly Pro Asn Arg Pro Thr Gly Pro Ser Ser Ser Thr Asp Leu Asn Thr
 435 440 445
 Tyr Val Asp Glu Ser Arg Ile Asp Arg Arg Ile His Pro Gly Asn Trp
 450 455 460
 Cys Asn Gln Ala Gly Ala Gly Leu Gly Glu Arg Pro Thr Val Asn Pro
 465 470 475 480
 Ala Pro Gly Val Asp Ala Tyr Val Trp Val Lys Pro Pro Gly Glu Ser
 485 490 495
 Asp Gly Ala Ser Glu Glu Ile Pro Asn Asp Gly Lys Gly Phe Asp
 500 505 510
 Arg Met Cys Asp Pro Thr Tyr Gln Gly Asn Ala Arg Asn Gly Asn Asn
 515 520 525
 Pro Ser Gly Ala Leu Pro Asn Ala Pro Ile Ser Gly His Trp Phe Ser
 530 535 540
 Ala Gln Phe Arg Glu Leu Leu Ala Asn Ala Tyr Pro Pro Leu
 545 550 555

<210> 15
 <211> 1293
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> construct based on *Thermobifida fusca*

<400> 15
 gccggtctca ccgccacagt caccaaagaa tcctcgtggg acaacggcta ctccgcgtcc 60
 gtcaccgtcc gcaacgacac ctcgagcacc gtctccagt gggaggtcg cctcaccctg 120
 cccggcggca ctacagtggc ccaggtgtgg aacgcccagc acaccagcag cggcaactcc 180
 cacacccctca cccggggttcc ctggaacagc accatcccgc cccggaggcac cgcctcttcc 240
 ggcttcatcg cttccggcag cggcgaaccc acccactgca ccatcaacgg cgccccctgc 300
 gacgaaggct cccgagccggg cggccccggc ggtcccccggaa ccccccctcccc cggccccggc 360
 acgcagcccg gcaccggcac cccggtcgag cggtacggca aagtccagggt ctgcggcacc 420
 cagctctgcg acgagcacgg caacccggc caactgcgcg gcatgagcac ccacggcatc 480
 cagtggttcg accactgcct gaccgacagc tcgctggacg ccctggccta cgactggaaag 540
 gccgacatca tccgcctgtc catgtacatc caggaagacg gctacgagac caacccgcgc 600
 ggcttcaccg accggatgca ccagctcatc gacatggcca cggcgcgcgg cctgtacgtg 660
 atcgtggact ggcacatcct caccccgggc gatccccact acaacctgga ccgggccaag 720
 accttcttcg cggaaatcgc ccagcgccac gccagcaaga ccaacgtgt ctacgagatc 780
 gccaacgaac ccaacggagt gagctgggcc tccatcaaga gctacgcccga agaggtcatc 840
 ccgggtatcc gccagcgcga ccccgactcg gtgatcatcg tgggcacccg cggctggcgt 900
 tcgctcggcg tctccgaagg ctccggccccc gccgagatcg cggccaaccc ggtcaacgccc 960
 tccaaacatca tgtacgcctt ccacttctac gcggcctcgc accgcgcacaa ctacctcaac 1020
 gcgctgcgtg aggcctccga gctgttcccg gtctcgtca ccgagttcg caccgagacc 1080
 tacaccgggt acggcgccaa cgacttccag atggccgacc gctacatcga cctgtatggcg 1140
 gaacggaaaga tcgggtggac caagtggAAC tactcggacg acttccgttc cggcgcggc 1200
 ttccagccgg gcacctgcgc gtccggcggc ccgtggagcg gttcgtcgct gaaggcgtcc 1260
 ggacagtggg tgccgagcaa gctccagtcc tga 1293

<210> 16
 <211> 430
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> construct based on *Thermobifida fusca*

<400> 16
 Ala Gly Leu Thr Ala Thr Val Thr Lys Glu Ser Ser Trp Asp Asn Gly
 1 5 10 15
 Tyr Ser Ala Ser Val Thr Val Arg Asn Asp Thr Ser Ser Thr Val Ser
 20 25 30
 Gln Trp Glu Val Val Leu Thr Leu Pro Gly Gly Thr Thr Val Ala Gln
 35 40 45
 Val Trp Asn Ala Gln His Thr Ser Ser Gly Asn Ser His Thr Phe Thr
 50 55 60
 Gly Val Ser Trp Asn Ser Thr Ile Pro Pro Gly Gly Thr Ala Ser Ser
 65 70 75 80
 Gly Phe Ile Ala Ser Gly Ser Gly Glu Pro Thr His Cys Thr Ile Asn
 85 90 95
 Gly Ala Pro Cys Asp Glu Gly Ser Glu Pro Gly Gly Pro Gly Pro
 100 105 110
 Gly Thr Pro Ser Pro Asp Pro Gly Thr Gln Pro Gly Thr Gly Thr Pro
 115 120 125
 Val Glu Arg Tyr Gly Lys Val Gln Val Cys Gly Thr Gln Leu Cys Asp
 130 135 140
 Glu His Gly Asn Pro Val Gln Leu Arg Gly Met Ser Thr His Gly Ile
 145 150 155 160
 Gln Trp Phe Asp His Cys Leu Thr Asp Ser Ser Leu Asp Ala Leu Ala
 165 170 175
 Tyr Asp Trp Lys Ala Asp Ile Ile Arg Leu Ser Met Tyr Ile Gln Glu
 180 185 190
 Asp Gly Tyr Glu Thr Asn Pro Arg Gly Phe Thr Asp Arg Met His Gln
 195 200 205
 Leu Ile Asp Met Ala Thr Ala Arg Gly Leu Tyr Val Ile Val Asp Trp
 210 215 220
 His Ile Leu Thr Pro Gly Asp Pro His Tyr Asn Leu Asp Arg Ala Lys
 225 230 235 240
 Thr Phe Phe Ala Glu Ile Ala Gln Arg His Ala Ser Lys Thr Asn Val
 245 250 255
 Leu Tyr Glu Ile Ala Asn Glu Pro Asn Gly Val Ser Trp Ala Ser Ile
 260 265 270
 Lys Ser Tyr Ala Glu Glu Val Ile Pro Val Ile Arg Gln Arg Asp Pro
 275 280 285
 Asp Ser Val Ile Ile Val Gly Thr Arg Gly Trp Ser Ser Leu Gly Val
 290 295 300
 Ser Glu Gly Ser Gly Pro Ala Glu Ile Ala Ala Asn Pro Val Asn Ala
 305 310 315 320
 Ser Asn Ile Met Tyr Ala Phe His Phe Tyr Ala Ala Ser His Arg Asp
 325 330 335
 Asn Tyr Leu Asn Ala Leu Arg Glu Ala Ser Glu Leu Phe Pro Val Phe
 340 345 350
 Val Thr Glu Phe Gly Thr Glu Thr Tyr Thr Gly Asp Gly Ala Asn Asp
 355 360 365
 Phe Gln Met Ala Asp Arg Tyr Ile Asp Leu Met Ala Glu Arg Lys Ile
 370 375 380
 Gly Trp Thr Lys Trp Asn Tyr Ser Asp Asp Phe Arg Ser Gly Ala Val
 385 390 395 400
 Phe Gln Pro Gly Thr Cys Ala Ser Gly Gly Pro Trp Ser Gly Ser Ser
 405 410 415
 Leu Lys Ala Ser Gly Gln Trp Val Arg Ser Lys Leu Gln Ser
 420 425 430

<211> 2656
<212> DNA
<213> Artificial Sequence

<220>
<223> fusion construct

<400> 17
atgtatcgga agttggccgt catctcgccc ttcttggcca cagctcggtc tcagtcggcc 60
tgcactctcc aatcgagac tcacccgcct ctgacatggc agaaatgctc gtctgggtgc 120
acttgcactc aacagacagg ctccgtggtc atcgacgcca actggcgctg gactcacgct 180
acgaacagca gcacgaactg ctacgatggc aacacttggc gctcgaccct atgtcctgac 240
aacgagacct gcgcaagaa ctgctgtctg gacgggtggc cctacgcgtc cacgtacgga 300
gttaccacga gcggtAACAG cctctccatt ggctttgtca cccagtctgc gcagaagaac 360
gttggcgctc gcctttacct tatggcgagc gacacgacct accaggaatt caccctgctt 420
ggcaacgagt tctcttcga tggatgtt tcgcagctgc cgtaagtgtac ttaccatgaa 480
cccctgacgt atcttcttgt gggctccag ctgactggcc aatttaaggt gcggcttgaa 540
cgaggctctc tacttcgtgt ccatggacgc ggatgggtgc gtgagcaagt atcccaccaa 600
caccgctggc gccaagtacg gcacggggta ctgtgacagc cagtgtcccc gcgatctgaa 660
gttcatcaat ggccaggcca acgttgaggg ctgggagccg tcataccaaca acgcaaacac 720
gggcatttggaa ggacacggaa gctgctgctc tgagatggat atctgggagg ccaactccat 780
ctcccgaggct cttacccccc acccttgac gactgtggc caggagatct gcgagggtga 840
tgggtggc ggaacttact ccgataaacag atatggcgcc acttgcgatc ccgatggctg 900
cgacttggAACAC ccataccgccc tggcaacac cagcttctac ggccctggct caagctttac 960
cctcgataacc accaagaaat tgaccgttgc caccctgttcc gagacgtcggt gcggccatcaa 1020
ccgataactat gtccagaatg ggttgcactt ccagcagccc aacgcccggc ttggtagtta 1080
ctctggcaac gagctcaacg atgattactg cacagctgag gaggcagaat tcggcgatc 1140
ctctttctca gacaaggggcg gctgactca gttcaagaag gtcacccctgc gcggcatgg 1200
tctggcatg agtctgtggg atgatgtgag tttgatggac aaacatgcgc gttgacaaag 1260
agtcaaggcg ctgactgaga tggatgttgc ctacgccaac atgctgtggc tggactccac 1320
ctacccgaca aacgagaccc cctccacacc cgggtggcgtc cgccggaaat gtcacccacc 1380
ctccgggtgc cctgctcagg tcgaatctca gtctccaaac gccaagggtca cttctccaa 1440
catcaagttt ggacccattt gcagcaccgg caacccttgc ggcggcaacc ctccggcg 1500
aaacccgctt ggcaccacca ccacccggcc cccagccact accactggaa gtcacccgg 1560
acctactagt aagcggggcg gggccggcta ttggcacacg agcggccggg agatcttgg 1620
cgcaacaac gtggccgtac ggatggccgg catcaactgg tttgggttcg aaacctgca 1680
ttacgtcggtc cacggctctt ggttgcacgcg ctaccggcggc atgctgtggc agataaaatgc 1740
gctcggtac aacacaatcc gggtggcgtc ctctgacgac attctcaaggc cgccggaccat 1800
gccgaacaggc atcaattttt accagatgaa tcaggacccctg cagggtctga cgtcccttgc 1860
ggtcatggac aaaatcggtc cgtacggccgg tcagatggc ctgcgcacatca ttcttgaccg 1920
ccaccggaccg gattgcagcg ggcagtcggc gctgtggtac acgagcagcg tctcgaggc 1980
tacgtggatt tccgacccgtc aagcgctggc gcagcgctac aaggggaaacc cgacgggtcg 2040
cggtttgtac ttgcacaacg agccgcatttgc cccggccgtc tggggctggc gcgttccgg 2100
catcgactgg cgattggccgg ccgagcgggc cgaaacgcg gtgttgcgttgc tgaatccaa 2160
cctgctcatt ttgttgcgttgc ggttgcagatg ctacaacggc gacttctact ggtggggcg 2220
caacctgcaa ggagccggcc agtacccggc cgttgcgttgc gtcggcaacc gcctgggtgt 2280
ctcgccgtac gactacggcga cgagcgctta cccggccgtc tgggttgcgtc atccgaccc 2340
ccccaaacaac atgcccggca tctggaccaa gaactggggc taccttgcgttca atcagaacat 2400
tgcacccggta tggctggcg aattcggtac gacactggc tccacgaccg accagacgtg 2460
gctgaagacg ctcgtccagt acctacggcc gaccggcgttgc gacgggtggc acagcttcca 2520
gtggacccctt tgggttgcgttgc accccggattt cggcgttgcgttgc gggggattt tcaaggatga 2580
ctggcagacg gtcgacacacg taaaagacgg ctatctcggttcccgatcaagt cgtcgattttt 2640
cgatcttgcgttgc ggctaa 2656

<210> 18
<211> 841
<212> PRT

<213> Artificial Sequence

<220>

<223> fusion construct

<400> 18

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg
1 5 10 15
Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr
20 25 30
Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser
35 40 45
Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser
50 55 60
Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp
65 70 75 80
Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala
85 90 95
Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe
100 105 110
Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met
115 120 125
Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe
130 135 140
Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala
145 150 155 160
Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro
165 170 175
Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
180 185 190
Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
195 200 205
Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly
210 215 220
Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu
225 230 235 240
Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu
245 250 255
Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr
260 265 270
Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr
275 280 285
Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys
290 295 300
Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr
305 310 315 320
Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly
325 330 335
Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu
340 345 350
Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln
355 360 365
Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp
370 375 380
Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr
385 390 395 400
Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr

405	410	415
Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys		
420	425	430
Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn		
435	440	445
Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr		
450	455	460
Thr Arg Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Ser		
465	470	475
Lys Arg Ala Gly Gly Tyr Trp His Thr Ser Gly Arg Glu Ile Leu		
485	490	495
Asp Ala Asn Asn Val Pro Val Arg Ile Ala Gly Ile Asn Trp Phe Gly		
500	505	510
Phe Glu Thr Cys Asn Tyr Val Val His Gly Leu Trp Ser Arg Asp Tyr		
515	520	525
Arg Ser Met Leu Asp Gln Ile Lys Ser Leu Gly Tyr Asn Thr Ile Arg		
530	535	540
Leu Pro Tyr Ser Asp Asp Ile Leu Lys Pro Gly Thr Met Pro Asn Ser		
545	550	555
Ile Asn Phe Tyr Gln Met Asn Gln Asp Leu Gln Gly Leu Thr Ser Leu		
565	570	575
Gln Val Met Asp Lys Ile Val Ala Tyr Ala Gly Gln Ile Gly Leu Arg		
580	585	590
Ile Ile Leu Asp Arg His Arg Pro Asp Cys Ser Gly Gln Ser Ala Leu		
595	600	605
Trp Tyr Thr Ser Ser Val Ser Glu Ala Thr Trp Ile Ser Asp Leu Gln		
610	615	620
Ala Leu Ala Gln Arg Tyr Lys Gly Asn Pro Thr Val Val Gly Phe Asp		
625	630	635
Leu His Asn Glu Pro His Asp Pro Ala Cys Trp Gly Cys Gly Asp Pro		
645	650	655
Ser Ile Asp Trp Arg Leu Ala Ala Glu Arg Ala Gly Asn Ala Val Leu		
660	665	670
Ser Val Asn Pro Asn Leu Leu Ile Phe Val Glu Gly Val Gln Ser Tyr		
675	680	685
Asn Gly Asp Ser Tyr Trp Trp Gly Gly Asn Leu Gln Gly Ala Gly Gln		
690	695	700
Tyr Pro Val Val Leu Asn Val Pro Asn Arg Leu Val Tyr Ser Ala His		
705	710	715
Asp Tyr Ala Thr Ser Val Tyr Pro Gln Thr Trp Phe Ser Asp Pro Thr		
725	730	735
Phe Pro Asn Asn Met Pro Gly Ile Trp Asn Lys Asn Trp Gly Tyr Leu		
740	745	750
Phe Asn Gln Asn Ile Ala Pro Val Trp Leu Gly Glu Phe Gly Thr Thr		
755	760	765
Leu Gln Ser Thr Thr Asp Gln Thr Trp Leu Lys Thr Leu Val Gln Tyr		
770	775	780
Leu Arg Pro Thr Ala Gln Tyr Gly Ala Asp Ser Phe Gln Trp Thr Phe		
785	790	795
Trp Ser Trp Asn Pro Asp Ser Gly Asp Thr Gly Gly Ile Leu Lys Asp		
805	810	815
Asp Trp Gln Thr Val Asp Thr Val Lys Asp Gly Tyr Leu Ala Pro Ile		
820	825	830
Lys Ser Ser Ile Phe Asp Pro Val Gly		
835	840	

<211> 10239
<212> DNA
<213> Artificial Sequence

<220>
<223> pTrex plasmid

<400> 19
aagcttaact agtacttctc gagctctgt a catgtccgg cgcgacgtac gcgtatcgat 60
ggcgccagct gcaggcgcc gcctgcagcc acttgcagtc ccgtggatt ctcacggta 120
atgttaggcct tttgttagggt aggaattgtc actcaagcac ccccaacctc cattacgcct 180
ccccataga gttcccaatc agtgagtcat ggcactgttc tcaaataagat tggggagaag 240
ttgacttccg cccagagctg aagggtcgcac aaccgcata tataagggtcg gcaacggca 300
aaaagcacgt ggctcaccga aaagcaagat gtttgcgatc taacatccag gaacctggat 360
acatccatca tcacgcacga ccactttgtat ctgctggtaa actcgtattc gccctaaacc 420
gaagtgacgt ggttaaatcta cacgtggcc ccttcggta tactgcgtgt gtcttctcta 480
ggtgcattc tttcccttc ctctagtgtt gaattgttg tggaggtc cgagctgtaa 540
ctacctctga atctctggag aatgggtggac taacgactac cgtgcacctg catcatgtat 600
ataatagtga tcctgagaag gggggtttgg agcaatgtgg gactttgatg gtcataaacc 660
aaagaacgaa gacgcctt ttgcaaagtt ttgttcggc tacggtaaag aactggatac 720
ttgtgtgtc ttctgtgtat tttgtggca acaagaggcc agagacaatc tattcaaaca 780
ccaagcttc tctttgagc tacaagaacc tgggggtat atatcttagat ttgtgaagtc 840
ggtaatcccg ctgtatagta atacgagtgc catctaaata ctccgaagct gctgcgaacc 900
cgggagaatcg agatgtgctg gaaagcttct agcgagcggc taaattagca tggaaaggcta 960
tgagaaattc tggagacggc ttgtgaatc atggcgatcc attcttcgac aagcaaagcg 1020
ttccgtcgca gtagcaggca ctcatcccg aaaaaactcg gagattccta agtagcgatg 1080
gaaccggaat aatataatag gcaatacatt gagttgcctc gacgggttgc atgcaggggt 1140
actgagctt gacataactg ttccgtaccc cacctcttca aaccccttgg cgtttccctg 1200
attcagcgta cccgtacaag tcgtaatcac tattaacca gactgaccgg acgtgtttt 1260
cccttcattt ggagaaataa tgcattgcg atgtgtatt tgcctgcattt accgactggg 1320
gctgtcgaa gcccgaatgt aggattgtt tccgaactt gctcgtagag gcatgttgc 1380
aatctgtgtc gggcaggaca cgccctcgaaat gttcacggca agggaaacca ccgatagcag 1440
tgtctagtag caacctgtaa agccgcaatg cagcatcaat ggaaaataca aaccaatggc 1500
taaaaagtaca taagttaatg cctaaagaag tcatatacca gcccgtataattt attgtacaat 1560
caagtggcta aacgtaccgt aatttgcacca cggcttgcgg ggttgcagaa gcaacggca 1620
agccccactt ccccacgtt gtttcttcac tcagtcattt ctcagctggt gatccccaa 1680
ttgggtcgct tgggtttcc ggtgaagtga aagaagacag aggttaagaat gtctgactcg 1740
gagcgtttg catacaacca agggcagtga tggaaagacag tggaaatgttg acattcaagg 1800
agtatttgc cagggatgtc tgagtgtatc gtgttgcgg gtttgcattt cgatacgc 1860
aatactgtat agtcacttct gatgaagtgg tccatattga aatgttaagtc ggcactgaac 1920
aggcaaaaga ttgagttgaa actgcctaag atctcggcc ctcggccctt cggcccttgg 1980
gtgtacatgt ttgtgcctcg ggcataatgc aagtgtggta ggatcgaaca cactgctgcc 2040
tttaccaagc agctgagggt atgtgtatgg caaatgttca gggccactg catggtttcg 2100
aatagaaaga gaagcttagc caagaacaat agccgataaa gatagcctca ttaaacggaa 2160
tgagcttaga ggcaaaagtca gcaatgtgtt atatataaaat gttcgaggc cgtgcctccc 2220
tcatgcttc cccatctact catcaactca gatcctccat gagacttgc caccatctt 2280
tgagggcacag aaacccaaata gtcacccgcg gactgcgcattt catgtatcg ggttggccg 2340
tcatctcgcc cttctggcc acagctcgatc ctcagctggc ctgcactctc caatcgaga 2400
ctcacccgccc tctgacatgg cagaaatgtt cgtctgggtt cacttgcactt caacagacag 2460
gctccgtggt catcgacgc aactggcgat ggcactcacgc tacgaacagc agcacaact 2520
gctacgatgg caacacttgg agctcgaccc tatgtcctga caacgagacc tggcgaa 2580
actgctgtct ggacgggtgc gcctacgcgtt ccacgtacgg agttaccacg agcggtaaca 2640
gcctctccat tggctttgtc acccagtctg cgcagaagaa ctttgcgttgc cgccttacc 2700
ttatggcgag cgacacgacc taccagaaat tcacccgttgc tggcaacgag ttctttcg 2760
atgttgcgtt ttcgcagctg cgcataatgc cttaccatca acccctgcacg tatcttcttgc 2820
tgggctccca gctgactggc caatataagg tgcggcttgc acggagctt ctacttcgtt 2880
tccatggacg cggatgggtt cgtgagcaag tatccacca acaccgcgtt cggccaagttac 2940

ggcacggggt	actgtgacag	ccagtgtccc	cgcgatctga	agttcatcaa	tggccaggcc	3000
aacgttggagg	gctgggagcc	gtcatccaaac	aacgcaaaaca	cgggcattgg	aggacacgga	3060
agctgctgct	ctgagatgga	tatctggag	gccaactcca	tctccgaggc	tcttacccccc	3120
cacccttgc	cgactgtcg	ccaggagatc	tgcgagggtg	atgggtgcgg	cggaacttac	3180
tccgataaca	gatatggcg	cacttgcgt	cccgtatggct	gcgactggaa	cccataccgc	3240
ctggcaaca	ccagcttcta	cggccctggc	tcaagcttta	ccctcgatac	caccaagaaa	3300
ttgaccgtt	tcacccagtt	cgagacgtcg	ggtgccatca	accgatacta	tgtccagaat	3360
ggcgtactt	tccagcagcc	caacgcccag	cttggtagtt	actctggcaa	cgagctcaac	3420
gatgattact	gcacagctga	ggaggcagaa	ttcggcggt	ccctttctc	agacaaggc	3480
ggcctgactc	agttcaagaa	ggctacctct	ggcggcatgg	ttctggcat	gagtctgtgg	3540
gatgatgtga	gtttgatgga	caaacatgcg	cgttgacaaa	gagtcaagca	gctgactgag	3600
atgttacagt	actacgcca	catgctgtgg	ctggactcca	cctaccgcac	aaacgagacc	3660
tcctccacac	ccggtgcgt	gchgccaagc	tgctccacca	gctccgggt	ccctgctcag	3720
gtcgaatctc	agtctccaa	cgccaaaggc	accttctcca	acatcaagtt	cggaccatt	3780
ggcagcaccg	gcaaccctag	cgcgccaaac	cctccggcg	gaaaccgc	ttgcaccacc	3840
accacccgc	gcccagccac	taccactgga	agctctccc	gacctactag	taagcggata	3900
aggcgcgcg	cgcgcagct	ccgtgcgaaa	gcctgacgc	ccggtagatt	cttggtgagc	3960
ccgtatcatg	acggcggcg	gagctacatg	gccccgggt	atttattttt	tttgtatcta	4020
cttctgaccc	ttttcaaata	tacggtcaac	tcatcttca	ctggagatgc	ggcctgctt	4080
gtattgcgt	gttgtcagct	tgccaaattt	ttggcttcga	aaacacaaaa	cgattccctta	4140
gtagccatgc	attttaagat	aacggaatag	aagaaagagg	aaattaaaaa	aaaaaaaaaa	4200
acaaacatcc	cgttcataac	ccgtagaatc	gccgctttc	gtgtatccca	gtaccagttt	4260
attttgaata	gctcgcccgc	ttggagagcat	cctgaatgca	agtaacaacc	gtagaggctg	4320
acacggcagg	tgttgcctagg	gagcgtcg	ttctacaagg	ccagacgtct	tcgcgggt	4380
tatatatgt	tgttgactg	caggctgctc	agcgacgaca	gtcaagttcg	ccctcgctc	4440
tttgtcaata	atcgcagtgg	ggaagccaca	ccgtgactcc	catctttag	taaagctctg	4500
tttgtgttta	tcagcaatac	acgtaattt	aactcgtag	catggggct	atagcttaat	4560
taccgtttac	cagtgcgcg	gttctgcagc	tttccttgc	ccgtaaaatt	cgcgaaagcc	4620
agccaatcac	cagctaggca	ccagctaaac	cctataatta	gtctcttac	aacaccatcc	4680
gctcccccgg	gatcaatgag	gagaatgagg	gggatgcggg	gctaaacaag	cctacataac	4740
cctcatgcca	actcccgat	tacactcg	gagccaaacat	cctgactata	agctaaca	4800
gaatgcctca	atccctggaa	gaactggccg	ctgataagcg	cgccgcctc	gcaaaaacca	4860
tccctgatga	atggaaagtc	cagacgctc	ctgcccga	cagcgttatt	gatttccaa	4920
agaaatcggg	gatccttca	gaggccgaac	tgaagatcac	agaggcctcc	gctgcagatc	4980
tttgttccaa	gctggcgcc	ggagagttga	cctcgggt	agttacgcta	gcattctgt	5040
aacgggcagc	aatcgcccag	cagttatgt	ggtcccctct	acctctcagg	gagatgtac	5100
aacgccac	tatggacta	tcaagctgc	gctggcttct	gtgcagacaa	actgcgc	5160
cgagttctc	cctgacgcgc	ctctcg	ggcaaggaa	ctcgatgaat	actacgcaaa	5220
gcacaagaga	cccgttggc	cactccatgg	cctccccc	tctctcaaa	accagctcg	5280
agtcaaggt	caccgttgc	cctaagtcgt	tagatgtcc	tttttgc	ctaacatatg	5340
ccaccaggc	tacgaaacat	caatggct	catctcatgg	ctaaacaagt	acgacgaa	5400
ggactcggtt	ctgacaacca	tgctccgaa	agccgggt	gtcttctac	tcaagac	5460
tgtcccgag	accctgatgg	tctgcgagac	agtcaacaac	atcatgggc	gcaccgt	5520
cccaacgc	aagaacttgt	cgtgcggcgg	cagttctgt	ggtaggggt	cgatcg	5580
gattcgttgt	ggcgtcatcg	gtttaggaac	ggatatcggt	ggctcgatt	gagtgc	5640
cgcgttcaac	ttccctgtac	gtctaaggcc	gagtcatgg	ccgctggcg	atgcaaa	5700
ggcgaacagc	atggagggtc	aggagacggt	gcacagcgtt	gtcggccg	ttacgc	5760
tgttgagggt	gagtccctcg	ccttccctt	cttttgc	tctataacc	gcctcc	5820
tcctccccc	ttgctttta	tactatatac	gagaccggca	gtcactgat	aagtatgtt	5880
gacccccc	tcttacccaa	atccgtc	ggtcaggagc	catggaaata	cgactcca	5940
gtcatccca	tgccctggcg	ccagtccg	tcggacatta	ttgcctccaa	gatcaaga	6000
ggcgggctca	atatcggt	ctacaacttc	gacggcaatg	tccttccaca	ccctcctatc	6060
ctgcgcggcg	tggaaaccac	cgtcggcg	ctcgccaaag	ccggtcacac	cgtgaccc	6120
tggacgccc	acaagcacg	tttcggcc	gatctcat	cccatatcta	cgcg	6180
ggcagcgc	acgtaatgc	cgatcat	gcatccggcg	agccggcgat	tccaaatatc	6240
aaagacctac	tgaacccgaa	catcaaagct	gttaacatga	acgagctctg	ggacacgc	6300
ctccagaagt	ggaattacca	gatggagat	tttgagaaat	ggcg	ggaggc	6360

gccccggaaagg aactggacgc catcatcgcg ccgattacgc ctaccgctgc ggtacggcat 6420
gaccagttcc ggtactatgg gtatgcctct gtgatcaacc tgctggatt cacgagcgtg 6480
gttggccgg ttacctttgc ggataagaac atcgataaga agaatgagag tttcaaggcg 6540
gttagtgagc ttgatgcctc cgtgcaggaa gagtatgatc cggaggcgta ccatgggca 6600
ccgggtgcag tgcaggttat cggacggaga ctcagtgaa agaggacgtt ggcgattgca 6660
gaggaagtgg ggaagttgct gggaaatgtg gtgactccat agctaataag tgcagatag 6720
caatttgcac aagaaatcaa taccagcaac tgtaataaag cgctgaagt accatgccat 6780
gctacgaaag agcagaaaaa aacctgccgt agaaccgaag agatatgaca cgcttccatc 6840
tctcaaagga agaatccctt cagggttgcg tttccagttc agacacgtat aacggcacaa 6900
gtgtctctca ccaaataatggg tataatctaa atgtgatcta aggatggaaa gcccagaatc 6960
taggcctatt aatattccgg agtatacgta gccggctaac gttacaacc ggtacctcta 7020
gaactatagc tagcatgcgc aaatttaaag cgctgatatc gatcgccgc agatccat 7080
atagggcccgg gttataatt acctcaggc gacgtcccat ggcattcga attcgtaatc 7140
atggtcatacg ctgtttccctg tgcggaaattt ttagtccgct acaattccac acaacatacg 7200
agccggaaagc ataaagtgtt aagcctgggg tgcctaattga gtgagctaac tcacattaat 7260
tgcgttgcgc tcactgcccgg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 7320
aatcggccaa cgcgcggggg gaggcggtt cggtattggg cgcttcccg cttcccgct 7380
caactgactcg ctgcgctcgg tcggtccggc gcccggagcg gtatcagctc actcaaaggc 7440
ggttaatacgg ttatccacag aatcaggggg taacgcagga aagaacatgt gagcaaaaagg 7500
ccagcaaaagc gccaggaacc gtaaaaaggc cgcgttgctg gcgttttcc ataggctccg 7560
ccccccctgac gagcatcaca aaaatcgacg ctcagtcg aggtggcgaa accccacagg 7620
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 7680
cctgcccctt accggatacc tgcgtccctt tctcccttcg ggaagcgtgg cgcttctca 7740
tagtcacgc tgcgttgcgt tcagttcggt gtaggtcggt cgctccaagc tgggtgtgt 7800
gcacgaaccc cccgttcagc ccgaccgctg cgccttaccc ggttaactatc gtctttagtc 7860
caacccggta agacacgact tgcgtccact ggcagcagcc actggtaaca ggattagcag 7920
agcgaggtat gtagggcggtg ctacagagtt ctgtaaatgg tggcctaact acggctacac 7980
tagaagaaca gtatggta tctgcgtct gctgaagcca gttacccctg gaaaaagagt 8040
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgttgc当地 8100
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttagtct tttctacggg 8160
gtctgacgct cagtgaaacg aaaactcactc ttaaggatt ttggcatga gattatcaa 8220
aaggatctt acctagatcc tttaaatataaaaatgaatg tttaaatcaa tctaaatgtat 8280
atatgatcaa acttggcttgc acagttacca atgcttaatc agtgaggcac ctatctcagc 8340
gatctgtcta ttgcgttcat ccatagttgc ctgactcccc gtcgttgcgta taactacgt 8400
acgggaggggc ttaccatctg gccccagtgc tgcaatgata cccgcgagacc cacgctcacc 8460
ggctccagat ttatcagcaa taaaccagcc agccggaaagg gcccggcgca gaagtggcc 8520
tgcaacttta tccgcctcca tccagtttat taattttgc cgggaagcta gagtaatgt 8580
ttcgccagtt aatagttgc gcaacgttgc tgccattgtc acaggcatcg tgggttcacg 8640
ctcgctgtt ggtatggctt cattcagctc cgggtcccaa cgatcaaggc gagttacatg 8700
atccccatg ttgtgc当地aaa aagcggttag ctccctcggt cctccgatcg ttgtcagaag 8760
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt 8820
catgccatcc gtaagatgtt tttctgtgac tgggtgatc tcaaccaagt cattctgaga 8880
atagtgtatg cggcgaccga gttgttgc cccggcgta atacggata ataccgc当地 8940
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc 9000
aaggatctt ccgtgttgc gatccagttc gatgtaaacc actcgtgcac ccaactgatc 9060
ttcagcatct ttactttca ccagcggttgc tgggtgatc aaaacaggaa ggcaaaaatgc 9120
cgcaaaaaag ggaataaggc cgacacggaa atgttgcata ctcataactct tccttttca 9180
atattattga agcatttac agggttattt tctcatgagc ggatacatat ttgtatgtat 9240
tttagaaaaat aaacaaatag ggggtcccgcc cacattccc cgaaaatgtc cacctgacgt 9300
ctaagaaacc attattatca tgacattaac ctataaaaat agggtatc cggggccctt 9360
tcgtctcgcc cggttcgggtt atgacgggtga aaacctctga cacatgcagc tccggagac 9420
ggtcacagct tgcgttgcag cggatgccgg gagcagacaa gcccgtcagg ggcgtcagc 9480
gggtgttggc ggggttgcggg gctggcttaa ctatgcggca tcaagtgatc ttgtactgt 9540
agtgcaccat aaaattgttac acgttataat tttgttaaaa ttgcgtttaa atttttgtt 9600
aatcagctca ttttttaacc aataggccga aatcgccaa atcccttata aatcaaaaaga 9660
atagcccgag atagggttgc gttgttgc agtttggaaac aagagtccac tattaaagaa 9720
cgtggactcc aacgtcaaag ggcggaaaac cgtctatcag ggcgttggcc cactacgtga 9780

accatcaccc	aatcaagtt	ttttgggtc	gaggtgccgt	aaagcactaa	atcgaaaccc	9840
taaagggagc	ccccgattta	gagcttacg	gggaaagccg	gcgaacgtgg	cgagaaaagga	9900
agggaaagaa	gcgaaaggag	cggcgctag	ggcgctggca	agtgtacgg	tcacgctgcg	9960
cgttaaccacc	acacccgccc	cgcattaatgc	gccgctacag	ggcgcgtact	atggttgctt	10020
tgacgtatgc	ggtgtgaaat	accgcacaga	tgcgtaagga	aaaaataccg	catcaggcgc	10080
cattcgccat	tcaggctgcg	caactgttgg	gaagggcgat	cggtgcgggc	ctcttcgcta	10140
ttacgcccagc	tggcgaaagg	gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	10200
ttttcccagt	cacgacgttg	taaaacgacg	gccagtgcc			10239
<210> 20						
<211> 12						
<212> DNA						
<213> Trichoderma reesei						
<400> 20						
actagtaagc gg						12
<210> 21						
<211> 41						
<212> DNA						
<213> Artificial Sequence						
<220>						
<223> primer						
<400> 21						
gcttatacta gtaagcgccg gggcgccggc tattggcaca c						41
<210> 22						
<211> 39						
<212> DNA						
<213> Artificial Sequence						
<220>						
<223> primer						
<400> 22						
gcttatggcg cgcccttagac aggatcgaaa atcgacgac						39
<210> 23						
<211> 43						
<212> DNA						
<213> Artificial Sequence						
<220>						
<223> primer						
<400> 23						
ctaagagaaa cgacccgtac atccagcggt tcctcacgat gta						43
<210> 24						
<211> 38						
<212> DNA						
<213> Artificial Sequence						
<220>						
<223> primer						

<400> 24
ttacccggat gggaaagagca tgccaaaatc ggcgttcg 38

<210> 25
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 25
ctaagagagc gacgactcag ccgtacacct ggagcaacgt ggc 43

<210> 26
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 26
ttacgatccg gacggcgcac caccaatgtc cccgtata 38

<210> 27
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 27
gcttatacta gtaagcgcgc cggctgctcg gtggactaca cg 42

<210> 28
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 28
gcttatggcg cgccttacag aggcgggtag gcgttgg 37

<210> 29
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 29

gcttatacta gtaagcgcgc cggcttcacc gccacagtca cc

42

<210> 30

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 30

gcttatggcg cgcctcagga ctggagcttg ctccgc

36n

21